P/61813.WOP 'A1

P.D. 1996 (5)
P. 223-227

A1

XP-002159253

TMN ARCHITECTURE FOR SDH NETWORKS USING IS–IS ROUTING PROTOCOL: DESIGN AND PERFORMANCES

Luca Ambrosoli*, Lothar Burger**, Mario Huterer*

ALCATEL TELECOM
TRANSMISSION SYSTEMS DIVISION

* Via Trento 30, 20059 Vimercate (Milan) – Italy ** Lorenzstraße 10, 70435 Stuttgart – Germany

ABSTRACT

This paper reflects selected results and guidelines which are extracted from an extensive study at Alcatel Telecom TSD on designing and optimising SDH TMN in IS–IS and ES–IS environment. A methodology for TMN design has been proposed. Typical questions like how to partition the network, how to set the numerous IS–IS configuration parameters in these NEs or how to achieve efficient reconfiguration capabilities are addressed. Performance issues are discussed and a numerical example is given which illustrates the behaviour of the protocol in an SDH network.

1 INTRODUCTION

To set up a Telecommunication Management Network (TMN), which is to be used for the management of Synchronous Digital Hierarchy (SDH) network, different link types can be used: e.g. Local Area Networks (LANs) to connect Network Elements (NEs) and the Operations Systems (OSs) within an office, X.25 links to connect the OS to remote NEs, or SDH Embedded Communication Channels (ECCs) to connect NEs which are located in different offices. The resulting TMN forms a packet switching network, the topology of which in general reflects the topology of the underlying SDH network and is thus characterized by a mixture of meshed, linear, star–shaped and ring–shaped structures.

Use of Intermediate System to Intermediate System (IS–IS) and End System to Intermediate System (ES–IS) routing protocols [1, 2] in such a TMN facilitates installation and operation due to "self learning" capability of these protocols and automatic network reconfiguration in case of failures. ES and IS systems exchange periodically "Hello" messages to verify about each other's existence. Link State messages (LSPs) are exchanged among ISs (level 1 and level 2) to communicate the current reachability of their neighbors and the cost to reach them. The entire "routing domain" is divided into "subdomains", also called "areas". Level 1 LSPs are exchanged among all IS NEs, i.e. all level 1 and level 2 systems which are in one area. Level 2 LSPs are exchanged among all level 2 NEs in the domain, thereby ensuring the connectivity in the entire routing domain. Each area has at least one level 2 IS. A level 2 IS is also a level 1 IS in its own area. LSPs are periodically generated or if there is a change in the network, e.g. change in the connectivity caused by a cable cut. Each system, i.e. each NE, maintains its own routing table obtained by the processing of LSPs. LSPs are spread out in the network using flooding principle. Each LSP is acknowledged (PSNP and CSNP messages). For details see [1,2].

An appropriate design of the TMN network for the given transmission topology is the key in order to achieve optimal behaviour of IS–IS protocol and at the same time optimize the network performance for the actual management traffic. TMN design may turn out to be difficult in case of complex underlying SDH network topologies. It may become complicated if foreign routing domains or single NEs which do not support ES–IS and IS–IS must be integrated. A well designed TMN from the routing point of view is characterized by minimum overhead traffic generated by IS–IS and ES–IS, reliable network topology and fast network reconfiguration upon failures. A bad design, however, can cause an impossibility to reconfigure the TMN, large delays in exchanging management information and large CPU loads in the NEs supporting IS–IS and ES–IS.

Chapter 2 is dedicated to SDH TMN design process methodology. Chapter 3 goes in more detail in some essential aspects when using IS–IS: hierarchical structuring of TMN using level 1 (L1) and level 2 (L2) routing mechanisms, major considerations in ES–IS and IS–IS configuration parameters which need to be provisioned in the NEs, as well as discusses a strategy on how to integrate NEs which are not supporting IS–IS and ES–IS and at the same time preserve reconfiguration capability of TMN. Chapter 4 reveals some illustrative performance aspects and refers to the related study in [3] which can be helpful in estimating the utilization of resources (CPU, memory, transmission bandwidth) needed for IS–IS and ES–IS protocols in typical network scenarios of interest (steady–state and sever failure conditions). Chapter 5 is a conclusion.

2 TMN DESIGN PROCESS

In the following we present our methodology for the systematic approach to design and optimize the TMN starting from the given SDH network topology. We propose to split the design process into several design steps (see Figure 1).

Generate starting TMN topology

As the starting point for the TMN design process we generate a TMN topology with minimum, sensible connectivity. To this end we first introduce intraoffice LAN within each office which contains multiple NEs. Thereafter, a minimum set of ECC based links is introduced to get a connected network.

Localize OS

In most cases the localization of the OSs will already be given by the organizational requirements of the customer. From a TMN design point of view the most preferable configuration is one in which the OS is co–located with multiple NEs which reside in the "centre" of the SDH network.

Partiti n the TMN network

In this step we partition the TMN into routing domains and/or subdomains.

There may be two types of NEs in a given network, those which support IS–IS and/or ES–IS routing protocols ("adaptive" routing NEs) and those which do not ("static" routing NEs). Because interworking between adaptive and static routing NEs may reduce the reconfiguration capability of the network, our target is to design a TMN topology which requires as less interworking between adaptive and static routing NEs as possible. Therefore we first try to partition the network into separate domains. Each of these domains is directly connected to the OS and consists of only one type of NEs, either adaptive or static routing NEs.

For the sake of simplicity we assume that all adaptive routing NEs are assigned to a single domain. Therefore, another design problem to be solved in this step is to partition the adaptive routing domain into subdomains. We focus in more detail on this in Chapter 3.1.

There may be network configurations in which it is not possible to assign all static routing NEs to separate domains e.g. if there are isolated static routing NEs which are surrounded by adaptive routing NEs. Such static routing NEs will be handled in a later design step.

Integrate OS

The problem to be solved in this design step is to define links between the OS and the static and adaptive routing domains. It must be decided to which NEs the OS is directly connected to. Furthermore, assuming that the OS supports adaptive routing, it must be decided to which adaptive routing subdomain (area) the OS is assigned to.

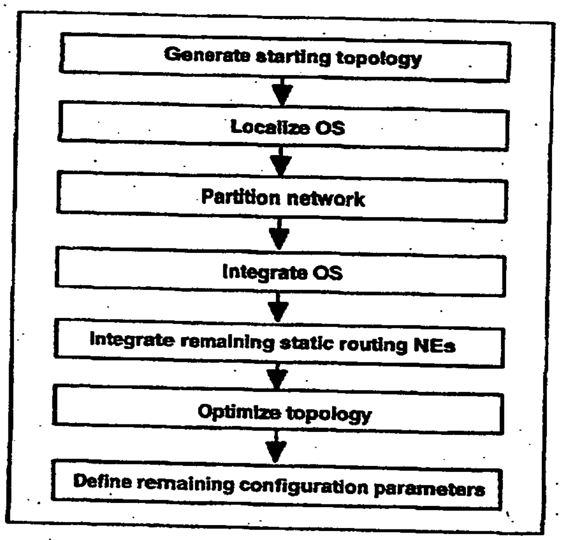

Figure 1. TMN design process

Integrate the remaining static routing NEs

In this step we focus on static routing NEs which could not be assigned to a separate static routing domain. They must be handled by introducing manual routing information in the adaptive routing NEs. Special care must be taken in this step to preserve reconfiguration capability of the TMN. We explain this in more detail in Chapter 3.2.

Optimize topology

As a starting point we assumed minimum connectivity of the TMN network. To reduce path lengths or to increase reliability, it might be necessary to introduce additional links. Topological ptimization should be performed taking into consideration both performance as well as reliability criteria. Of course, additional links introduce additional cost (e.g. X.25 links) which needs to be considered.

To allow for satisfactory response times, path length between an OS and any NE should not be excessive as, of course, routing at each hop introduces certain delay. Furthermore care should be taken that certain links or NEs do not become bottlenecks, because they have to carry traffic i.e. ES–IS and IS–IS protocol traffic as well as the actual management traffic originated by or destined to a number of other NEs.

As a simple measure for the reliability of the communication path between an OS and an NE the connectivity measure can be used. The connectivity is defined as the minimum number of link or node failures which disconnect an NE from the OS.

Topological optimization is strongly related to network partitioning and integration of the OS. Therefore it might be necessary to go back to previous design steps to revise outcomes of these steps until a satisfactory solution is identified.

In general, when optimizing the behaviour of IS–IS, the aim should be to meet the following objectives:

- Reduce the overhead traffic originated by the IS–IS protocol. The benefit of this is in an increase in the available routing capacity and bandwith for the actual management traffic.
- Reduce the amount of memory required by each NE to store the databases associated with IS–IS protocol and the routing tables generated by the "decision process". Memory limitation is particularly important for those transmission NEs which are required to act as ISs, i.e. to perform routing, but for which the memory cost plays an important factor (e.g. Add Drop Multiplexers, Line Systems or Regenerators).
- Reduce the processing overhead for the management of IS–IS databases used in the "update process" and "decision process". This is related to sizing of subdomains (number of ISs within the same subdomain).
- Reduce the CPU load which is related to the choice of IS–IS configuration parameters. One principal configuration parameter is the choice between L1 or L2 IS (L2 systems require more resources both in memory and processing power).

At the end of this design step, we can finally fix the role of each adaptive routing NE. NEs which have to take routing decisions must become intermediate systems. All other NEs should be made end systems to reduce the amount of protocol traffic. ISs which have a link to a foreign subdomain must become L2 ISs, all other ISs can be made L1 ISs.

Define the remaining configuration parameters

In the final design step we define the remaining configuration parameters of the OS and each NE. The amount of configuration parameters to be defined for a system mainly depends on the number and type of links attached to the system and the role of the system in the network. Typical configuration parameters are for example manualAreaAddresses, manualL2OnlyMode, circuitType, maximumPathSplits or security related parameters. Refer also to the related discussion in Chapter 3.3.

3 IS–IS NETWORK DESIGN GUIDELINES

In this chapter we address specifically the network partitioning task as well as NE IS–IS configuration and handling of non–adaptive routing NEs.

3.1 Network partitioning

TMN network planner shall split the routing domain into subdomains exploiting the benefits of hierarchical routing in IS–IS protocol.

TMN network partitioning requires identification of L1 subdomains. In many cases the structure of the TMN network and its consequent partitioning arises naturally from the topological structure of the underlying SDH network. An example for L1 subdomains could be: a digital cross–connect (DXC) and its attached ADM–rings, a large single ADM ring, a higher level ADM ring and its attached lower level rings, etc.

In other cases, however, we may face a meshed network composed of many homogeneous NEs and no natural partitioning can be seen at first glance. Nevertheless, it is necessary to identify groups of NEs (usually geographically close to each other) and identify a level 2 backbone. Use of commercial routers/bridges and leased lines (possibly extracted from the transmission network itself) or X.25 Virtual Circuits can be considered when designing the level 2 backbone.

After having identified L1 subdomains, L2 ISs are selected by making each NE, which has links to foreign areas, a L2 IS.

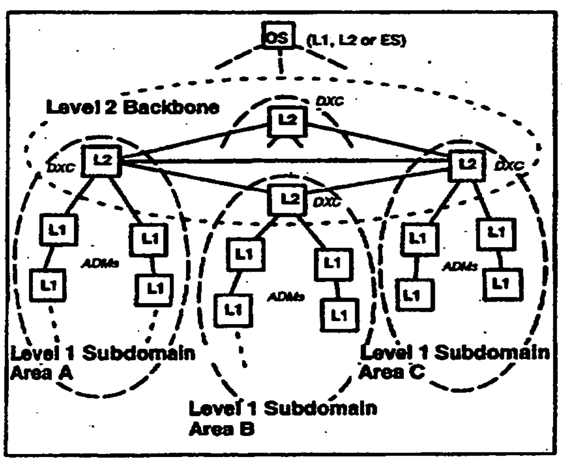

Figure 2. Partitioning of a routing domain

In the following we propose guidelines which should be taken into account when partitioning the TMN network into routing subdomains:

a) Non overlapping of level 1 subdomains.
Each NE shall belong to only one level 1 subdomain.

b) Connectivity within a L1 subdomain.
In order to guarantee network reconfigurability and to minimize the use of the IS–IS "partition repair" functionality, the connectivity within the L1 routing subdomain shall be higher than connectivity to outside subdomains.
A single link failure within a level 1 routing subdomain should not cause its partition into two subdomains and thereby force the use of the partition repair function.

c) Size of level 1 subdomains.
The size of a level 1 subdomain in terms of a number of NEs is a very critical issue. Basically many aspects shall be taken into account: performance limitations, acceptable round–trip delay, throughput of the network, throughput of bottleneck NEs (typically the gateway NE), estimated CPU load devoted to the IS–IS process, memory required by IS–IS databases, and so on.

d) Number of different level 1 subdomains.
The number of level 1 subdomains within the TMN network, together with their size, has a big impact on the IS–IS related memory and CPU load requirements. The bigger the number of areas the smaller the number of NEs per area, consequently, the lower the memory required to store level 1 intermediate system information (note: both L1 and L2 ISs need to be taken into account, as L2 systems are L1 systems in their own area). However, the more areas the more L2 systems, hence, the bigger the memory required to store L2 information (only L2 ISs).

e) Size of level 2 subdomains.
In order to take full advantage of hierarchical partitioning in terms of reduction of overhead, a number of L2 ISs with respect to the total number of NEs in the area should be small.

f) Connectivity of level 2 subdomains.
It is mandatory for L2 ISs to form a connected network. In fact a partition in the level 2 network causes major problems in the global connectivity of the network. Occasionally, redundant links via an external network (e.g. X.25), need to be introduced to guarantee the global connectivity of all the L2 ISs of the TMN network.

Nevertheless, despite any possible guideline, it shall be recognized that practical experience is still necessary to design an efficient TMN.

An example of a resulting partitioned SDH TMN with allocated IS systems is shown in Figure 2: DXCs as L2 systems forming a backbone routing network and ADMs in the subtended rings as L1 subsystems.

Figure 3 shows an example on how an operator may try to obtain a more reliable TMN, however while doing so, obtains an improper configuration. Let us consider the network topology in this figure. The two ADM rings with their gateway NEs[1] are attached to the same 802.3 LAN. ADM rings are configured as two different areas. In addition, to increase reliability, the rings are connected through a DCC channel on an existing Synchronous Transport Module (STM) link between nodes D and M (DL2 link). Moreover, to further increase reliability and reduce path lengths, nodes A and D are interconnected via additional X.25 link (link DL1).

Now, assume that the X.25 link fails. This causes the partition of the level 2 backbone and the impossibility to exchange L2 PDUs between nodes D and M and nodes A and G. In such situation, even though it may seem at a first sight that the connectivity between the OS and the NEs is guaranteed, due to the IS–IS mechanism, NEs C, D, E, L, N and M cannot reach the OS anymore. These NEs refer to the nearest L2 ISs (that is the NEs D and M) to route management packets towards the OS, but these L2 ISs have no knowledge on how

1) In this text a "gateway NE" is an NE which links an ECC subnetwork via an 802.3 LAN to an OS.

to reach the routing subdomain of the OS. Therefore, all packets addressed to the OS will get lost.

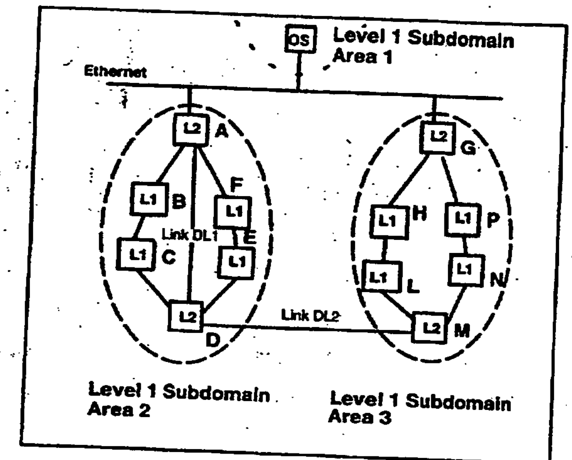

Figure 3. Example of a badly configured network

A right design for this network would require either introduction of an additional link connecting NE D to another L2 router (in order to avoid the partition of the level 2 backbone as a consequence of a single link failure), or merging of the level 1 subdomains into one, or disabling of the DCC channel on the STM DL2 link.

As a second scenario which can be examined on the same topology let us consider a failure of the DL2 link. Due to the fact that NE M is not connected to any outside routing subdomain anymore, it looses its role of the "reference L2 IS" for its subdomain. Therefore, all the NEs of the second ring would still be able to talk to the OS referring to the NE G as the reference L2 IS. The same applies in case of a simultaneous failure of links DL1 and DL2.

We can conclude that particular care must be taken in order to avoid a situation in which a single link failure generates a partition into subdomains which are composed of two or more L2 ISs belonging to two or more separate level 1 subdomains.

3.2 Integrating with non IS–IS or ES–IS NEs

It can happen that the TMN planner is obliged to deal with routing subdomains, or single NEs, which are not supporting IS–IS nor ES–IS protocol. Interworking between adaptive and static routers may reduce the reconfiguration capability of the TMN to a high degree. Therefore the target should be to design a topology which leads to separate traffic flows for the adaptive and static domains toward the OS(s). Hence, whenever possible, it is recommended to create entire routing subdomains containing only static routers and to interconnect these subdomains to the IS–IS network via Reachable Address Prefixes. Of course, it must be taken into account that static routing information is not affected by network failures (unless in some particular cases of point–to–point link failures) and, therefore, network reconfiguration cannot always take place.

As an example, let us consider the interconnection of a static domain composed of three SDH regenerators with an adaptive routing domain, a ring of ADMs, in Figure 4. The definition of a Reachable Address Prefixes on both NE 1 and NE 5 guarantees correct interworking within the IS–IS network.

The Multiplexer Section DCC guarantees direct interconnection between NE1 and NE5. A failure of either DL1 or DL4 Regenerator Section DCCs is detected by the IS–IS protocol on NE1 and NE5, and therefore, network reconfiguration will occur ("manual routing" entries are automatically disabled as soon as Data Link Layer disconnection is detected on the relevant interfaces). Nevertheless, a failure of the DL2 link cannot be recognized, as it is nested within the static routing domain. Therefore, no network reconfiguration can take place and NE1 will keep on sending packets addressed to NE3 through DL1 (even if NE3 would be reachable through NE5 and NE4), causing the unreachability of NE3.

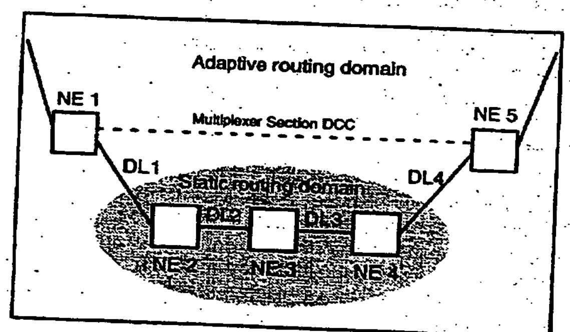

Figure 4. Interworking of static and adaptive routing domains

Concerning the failures of internal links of the adaptive routing domain, we face the same problem: static routing domain is not able to learn about failures inside adaptive routing domain. As a consequence, communication to the OS might be broken as well.

As a conclusion, to avoid some non–standard solutions which may bring workarounds in some specific cases, it is absolutely advisable to try to place static and adaptive routers into separate domains.

3.3 Configuration of IS–IS parameters

A sophisticated design of the network implies correct setting of IS and ES configuration parameters. Apart from the basic choice of a hierarchical level of the NE, i.e. level 1 versus level 2, which significantly affects the required memory and CPU power, other essential parameters are:

- Definition of external domain interfaces. This configuration avoids the transmission of IS–IS PDUs toward those subnetworks with which IS–IS interworking is not required or subdomains not implementing IS–IS protocol.
- Setting of Level2Only parameter on relevant interfaces in order to avoid transmission and reception of level 1 PDUs toward external areas.
- Setting of default metrics. In doing this, the operator can prioritize routing of packets over some subnetworks with respect to others (typically Ethernet with respect to ECC links). Despite the protocol offers four different routing metrics, using only the default metric we consider enough to tune the network.
- Setting of priority assigned to NEs to elect the designated Intermediate System within a broadcast subnetwork. In this way, similarly to the previous point, we can "force" systems with larger CPU capacity (frequently an OS, or NEs with larger routing processing capabilities – usually with a dedicated routing processor – r simply NEs less loaded by routing activity)

to work as "designated" Intermediate Systems over a pseudo-node, which leads to a more adequate distribution of the CPU utilizations in a broadcast network.

4 PERFORMANCE ISSUES

In this chapter we try to exemplify some performance issues of IS–IS protocol. It is well known that the amount of protocol traffic can be fairly extensive. The usual questions are how much protocol traffic is added on the TMN, what is the TMN reconfiguration time after link failures (e.g. fiber cut), if the amount of protocol traffic is high how is then the actual management traffic affected by the introduction of IS–IS, etc.? The answers to the above questions are, of course, linked to a number of parameters such as the underlying network topology, choice of NEs as L1 or L2 ISs, setting of IS–IS parameters (like "Hello" timers, ...) etc. The network topology together with the choice of hierarchical IS level basically determines the number of "neighbors" with which ISs need to exchange protocol messages. During the regular, i.e. steady–state network condition where no change in the network occurs, the "Hello" and Link State PDU (LSP) messages are exchanged in periodical intervals according their respective timer settings. In the situation where network topology change or change in the network configuration parameters occurs, LSP messages are generated, flooded throughout the network and the network reconfiguration process starts.

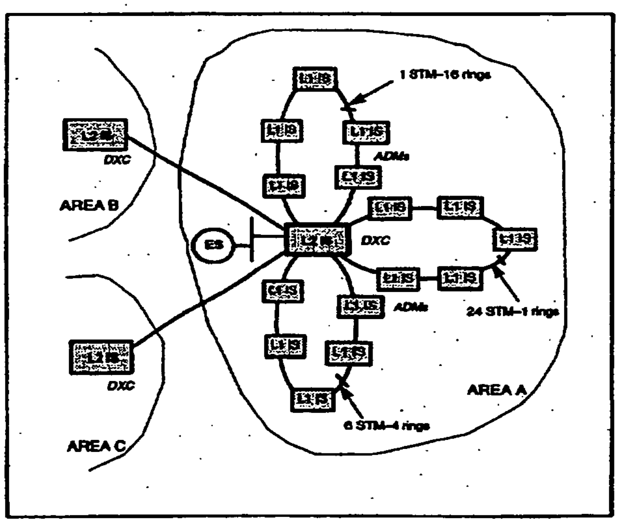

Figure 5. SDH network example: a 4/1 DXC with subtended ADM rings in a regional network

For an illustrative numerical example we refer the reader to study in [3], in which network of Figure 5 is analysed in detail. The results reveal that the average Hello traffic handled at the router of the gateway NE (the 4/1 DXC) turns out to be around 157 kb/s and the contribution of the LSP traffic amounts to maximum 65 kb/s, acknowledgement traffic equals to 3 kb/s. In this particular case the Hello message size is assumed 511 bytes (LAP–D message size is 512 bytes), L1 LSPs 773 bytes, L2 LSPs 97 bytes, acknowledgements (PSNPs) 35 bytes; Hello messages are exchanged every 3 seconds, LSPs every 900 seconds (if no change in the network occurs). Concerning the sizes of Hello messages, LSPs and acknowledgements reader may refer to [1] and compute the protocol traffic load for his particular network.

The very high values in the above example are due to the assumption that, in this particular network, the 4/1 DXC is directly connected to 31 ADM rings, which means that there are actually 62 L1 IS neighbors. This is, of course, a quite exagerated case, but it can serve to illustrate a potential scenario for a protocol traffic load of a node with a very high TMN connectivity.

In addition to the detailed IS–IS traffic analysis, the related study in [3] offers the possibility to understand the relation of the protocol traffic in its magnitude and character with respect to all other traffic, i.e. actual management traffic, in this network. Reader can refer to this study and use the methodology for a rough traffic analysis for his particular network of interest.

During the course of the study of SDH TMN design and optimization we have added to the numerical traffic studies as the one in [3], a number of traffic simulation studies in which for different typical network topologies the behaviour of ES–IS and IS–IS protocol traffic has been examined in the presence of actual management traffic loads. We gained further practical experiences from measurements during network validations in our test plants as well as actual network deployments. Our experience shows that, generally speaking, despite of the fact that largest number of messages in the networks are the actual IS–IS protocol messages, we have obtained no overload situations due to the protocol in our routers and the network reconfiguration after failures (like a fiber cut) has been fairly quick. The actual value varies from case to case, however, in a well designed TMN the network reconfiguration time can be typically kept in the range of one second, or even smaller.

5 CONCLUSIONS

The use of standardized routing protocols guarantees interoperability in a multivendor environment. Therefore, there is an increasing interest in using these protocols within SDH TMN. Our experiences show that the use of this protocol in SDH TMN has been now very well understood from its implementation to its configuration and performance optimization in the actual network deployments. We have also addressed the issue of co–existence of non IS–IS and IS–IS NEs in the same TMN and proposed example solutions. Of course, the use of both ES–IS and IS–IS is not a completely straight forward task. A number of issues must be considered during the TMN design process in order to optimize its performance in each network topology. Hence, in general, each topology is yet a new design assignment for the TMN designer. However, experience can be quickly obtained and bad designs, to which we tried to point to in this paper, minimized. Our intention is that this paper contributes to the clarity on how to use these protocols successfully in the SDH TMN.

REFERENCES

[1] ISO/IEC IS 10589: Intermediate System to Intermediate System Intra–Domain Routing Information Exchange Protocol, January 1992.

[2] ISO/IEC ES 9542: End System to Intermediate System Routing Exchange Protocol, 1988.

[3] SDH management: Traffic Analysis in a CMISE Environment, ICCT '94, Shanghai, June 1994.

THIS PAGE BLANK (USPTO)